

AD-A203 097

REPORT DOCUMENTATION PAGE

1a. SECURITY CLASSIFICATION AUTHORITY NA		1b. RESTRICTIVE MARKINGS NATL FILE COPY	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA		3. DISTRIBUTION/AVAILABILITY OF REPORT Distribution unlimited	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Mass. General Hospital/Harvard Med. School		5. MONITORING ORGANIZATION REPORT NUMBER(S) NA	
6a. NAME OF PERFORMING ORGANIZATION Massachusetts General Hospital	6b. OFFICE SYMBOL (If applicable) NA	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research	
6c. ADDRESS (City, State, and ZIP Code) Massachusetts General Hospital Jackson 13, Boston MA 02114		7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington VA 22217-5000	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research	8b. OFFICE SYMBOL (If applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER n00014-86-K-0116	
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington VA 22217-5000		10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61153N	TASK NO. RR04106
11. TITLE (Include Security Classification) (U) Computational Approaches to Protein Structure Design: beta-Bellin			
12. PERSONAL AUTHOR(S) Novotny Jiri			
13a. TYPE OF REPORT FINAL	13b. TIME COVERED FROM 2/1/86 TO 30/89	14. DATE OF REPORT (Year, Month, Day) 1/23/89	15. PAGE COUNT 17
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES FIELD GROUP SUB-GROUP 08		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Protein engineering, 3D structure, Molecular dynamics	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Based on the known principles that govern protein folding of polypeptide chains into anti-parallel B-sheet structures, a small protein molecule, B-bellin, has been designed by Drs. Jane and David Richardson. In the current project, computer methods were used to evaluate the structure of B-bellin models and, more generally, to answer fundamental questions about the relations among molecular geometry, atomic potential energy, and static stability of B-sheeted proteins. Using CONGEN, a program for conformational search, macromolecular energy, minimization and dynamics calculations, the stereochemistry, molecular surface, B-sheet twist and properties of B-bellin were analyzed. To elucidate the role of backbone and side chains in turn twist, molecular dynamics simulations were done on beta-bellin and its variants where the prolines 8, 16, and 24 were replaced either with glycines or with D-prolines. The simulations indicated that the most frequent conformation of the turns, in the absence of hydrogen bonding, was the left-handed type I' twist.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION U	
22a. NAME OF RESPONSIBLE INDIVIDUAL M. Marron		22b. TELEPHONE (Include Area Code) (609) 683-6209	22c. OFFICE SYMBOL ONR

CONFIDENTIAL

#19. ABSTRACT (continuation from page 1)

L-proline in position 2 of the turn, aspartate at position 3, and N-O hydrogen bonds between position 1-4 reversed this tendency and induced the right-handed, type I twist instead. Our results make it possible, in principle, to design reverse turns of predetermined stereochemistry in protein structures. Keywords: Protein engineering, (RW)

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
and/or	
Dist	
A-1	

89 1 27 018

FINAL REPORT:
"COMPUTATIONAL APPROACHES TO PROTEIN STRUCTURE DESIGN: β -BELLIN".
February 1986 - January 1989

Summary of Project Goals.

Based on the known principles that govern protein folding of polypeptide chains into antiparallel β -sheet structures, a small protein molecule, β -bellin, has been designed by Drs. Jane and David Richardson. Amino acid sequences of several β -bellin variants have been synthesized by Dr. Bruce Erickson. In the current project, computer methods were used to evaluate the structure of β -bellin models and, more generally, to answer fundamental questions about the relations among molecular geometry, atomic potential energy, and static stability of β -sheeted proteins. Using CONGEN, a program for conformational search, macromolecular energy, minimization and dynamics calculations, the stereochemistry, molecular surface, β -sheet twist and properties of β -bellin were analyzed and compared with those of other typical β -sheeted proteins such as immunoglobulins, prealbumin or superoxide dismutase. The project was particularly directed towards improving the unsatisfactory features of solvent-accessible surface of the β -bellin molecule. Quenched dynamics simulations were used to improve on surface smoothness and side chain interactions.

Summary of Accomplishments

(1) Preliminary characterization of β -bellin model.

Coordinates of the model of β -bellin were obtained from Jane Richardson and Richard Feldmann. Potential energy of the model was examined by the program CHARMM, version 16. The energy of the model (β -bellin dimer with 588 atoms) was comparable to that of an unrefined crystallographic structure (≈ 1 kJ/atom). The geometry of the model was not optimal but readily improved upon energy minimization.

A disulfide-bonded version of β -bellin was constructed next by substituting side chains of Val 21 for cysteines (sulphur atoms were taken to be identical to the CG1 carbons of the Val residue). The construction did not introduce any large strain into the structure. After 200 cycles of the ABNR minimization protocol, the total energy of the structure improved to ≈ -1700 kJ. Although the C β -S-S-C β torsional angle was not optimal, it seemed clear that manual adjustment of this torsion to, e.g., 90° followed by additional energy minimization would produce a flawless disulfide.

The Chothia-Janin-Teller formula for the exposed surface of a single-domain protein gives the expected accessible surface for β -bellin 39.98 nm^2 . The surface obtained using the Lee & Richards (1971) algorithm on the (unminimized) model was 45.13 nm^2 , i.e. 13% larger. This difference may be indicative of an incorrect fold. On the other hand, the ratio of nonpolar/polar surface side chain atoms for β -bellin was 1.2 while it was about 1.7 for native proteins and above 2.0

for incorrectly folded models.

The last property of the β -bellin model that was systematically examined was the β -sheet twist. Straight lines were fitted into the β -strand backbones by a least-square procedure and the angles between successive β -strand axes were measured. All the strand-strand angles were exceptionally low (4° ; 2.4° ; 5.6°) and remained so after minimization. The unusual flatness of β -bellin barrel can be visualized by least-squares fitting of a twisted hyperboloid into the β -strands. Such surfaces, when fitted to common β -barrels, have semiaxes of elliptical cross-sections of about 0.9 nm and 0.6 nm, respectively, with helical pitch in between 8-9 nm. The β -bellin hyperboloid, on the other hand, was characterized by semiaxes lengths 2.5 nm and 0.5 nm, and its helical pitch was very high (61.3 nm) indicating virtually no twist.

(2) Molecular representation of the BAB linker.

One of the more successful β -bellin preparations contained a bivalent linker, BAB, which tethers the two C-termini of the β -bellin dimer. The linker is a complex organic molecule combining aryl derivatives with motifs of peptide chemistry. In order to represent correctly the structure and stability of the BAB-linked β -bellin dimer, a "Residue Topology File" and "Parameter File" of the BAB molecule were prepared, the BAB molecule representation was generated in the CONGEN program, and its stereochemistry and energetics was analyzed. Dr. R. E. Bruccoleri participated in BAB molecule construction.

(3) Analysis of backbone shapes, backbone twist and side chain interactions in β -sheeted protein domains.

One of the approaches to the problem of protein folding has been to study protein shapes. Single domains built of antiparallel β -strands were alternatively described as twisted prisms (C. Chothia), elliptical barrels (J. Richardson) and twisted hyperboloids (J. Novotny, R. Bruccoleri & J. Newell). We have completed a mathematical least-squares-fitting study of higher-power curved geometrical surfaces into polypeptide backbones of 8 different β -barrels. Current models of β -barrels involve the supposition of a regular average right-handed twist of the β -strands; we have found that large deviations from the average β -strand twist values occur in positions of polar side chains. In superoxide dismutase, the exceptional backbone twist values (exceeding $\theta = +20^\circ \pm 20^\circ$) are so common that the supposition of a regular backbone twist has to be abandoned. A profile of side chain energies is dominated by large fluctuations, showing the striking anisotropy of atomic force distribution throughout the structure. The regular right-handed twist seems to be characteristic of the non-polar, spherically symmetric van der Waals forces among side chain atoms, whereas hydrogen bonding, with its strict directionality, tends to disrupt this pattern. The picture of β -barrel energetics which emerged from the analysis is an unexpected one. Regular β -barrels can be constructed from coiled and irregular backbones, and deviations from the ideal, regularly twisted β -sheet geometry are quite common. The importance of geometrical models, such as the twisted hyperboloid, thus becomes apparent: an abstract mathematical model facilitates analysis of β -barrel properties and helps to bridge the gap between the geometrical and physical attributes of the structures.

(4) Electrically charged molecular surface of β -bellin.

Following a suggestion of Dr. Jane Richardson, electrically charged solvent accessible surfaces were computed in a series of proteins, and compared with the percentage of electrically charged surfaces in various β -bellins. The results showed that the soluble proteins to have at least 10% of their solvent-exposed surface charged. The original β -bellin model had 7.3% charged surface, whereas the BAB-linked β -bellin, with dramatically improved solubility, has 11.2% charged surface.

(5) Quenched molecular dynamics simulations.

Quenched dynamics consists of molecular dynamics simulation at a predetermined temperature, followed by a gradual cooling of the structure. It is presently the most powerful energy minimization protocol available for computer manipulation of protein structures.

In the first run, the heating phase of the simulation involved 3 ps of 10° K temperature increments every 100 fs spanning the temperature range from 0° K to 300° K. During each temperature increase the velocities of atoms were reassigned according to a Gaussian distribution. Ten ps of simulation were then carried out, followed by 3 ps of cooling; the cooling represented a reversal of the heating protocol, *i.e.* 10° K temperature decrements every 100 fs from 300° K to 0° K. The second run involved 5 ps heating phase, from 0° K to 500° K, 10 ps of dynamical simulation at 500° K, and 5 ps of cooling down to 0° K. Figure 1 shows the time course of total energy (light line) and temperature (heavy line) changes in the simulations. Several interesting points emerged from the analysis of the completed dynamics runs.

- (i). Relatively large movements were apparent throughout the whole structure, including the BAB linker. At higher temperature (500° K) the N-terminal β -strand showed tendency to unfold away from the rest of the structure.
- (ii). Significant rearrangements were visible at the tight turns which include proline residues. Here, the backbone twist angle θ changed quite dramatically, compared to the starting β -bellin structure.
- (iii). Perhaps most importantly, the relatively flat β -bellin sheet acquired a distinct right-handed sheet twist on quenched dynamics. This was in contrast to the left-handed sheet twist found in all the naturally occurring β -sheets. It seems that the driving force of sheet twisting originated from the proline-containing turns. Thus, analysis of energetics and geometry of the quenched β -bellins concentrated on the possible role of β -turns in β -sheet architecture.

(6) Molecular dynamics simulations to explore β -turn conformations.

In order to get insight into the role of side chains in β -turn conformation, dynamical simulations were run not only on the original β -bellin model but also on its sequence variants where proline residues were replaced with either glycines or D-prolines. The results of the simulations demonstrated convincingly the relationship between β -turn twist, side chain chirality, and the hydrogen bonding pattern inside the turn. The turns 7-8-9-10 and 15-16-17-18 of the original β -bellin coordinates had the canonical hydrogen bonding pattern (see also Fig. 2) whereas the turn 23-24-25-26 had an abnormal hydrogen bond between the residues 24 and

26, more reminiscent of γ -turns (see also Fig. 2). This difference seemed to predetermine all the 24-25 turns to become type I' (lefthanded, see Fig. 3) while the other turns acquire either the righthanded type I (Fig. 4) or the type I' twist.

The greater backbone flexibility of 8,16,24-Gly- β -bellin was reflected in larger fluctuations of turn twist values. Nevertheless, the trends observed in β -bellin were reproduced in the Gly- β -bellin. That is, the type I twist was stabilized by the side chain carboxylate hydrogen bonding, and type I' turns tended to dispense with the unusual Phe-Pro N \rightarrow O hydrogen bond as their twist became more negative.

Introduction of the D-proline into β -bellin turns was accompanied by negative (type I'-like) twists. Turns which, upon dynamics, maintained the backbone-backbone hydrogen bonds, or acquired the backbone-side chain carboxylate hydrogen bonds, tended to acquire more positive twists. The type I' twist seemed to be stable even when backbone-backbone hydrogen bonds vanished.

(7) Molecular dynamics simulations with implicit solvent effects.

During the last part of the funding period, the quenched molecular dynamics simulations were repeated under conditions that simulated more realistically solvent effects. That is, (i) the non-bonded interactions were evaluated to infinity; (ii) the AMBER/CHARMM 19 parametrization, with larger van der Waals radii and larger partial atomic charges (simulating polarization effects) was employed; (iii) and the electrostatic potential function had the form $E_{el} = (Q_1 Q_2) / (16\pi R^2)$, where the "dielectric constant", $D = 4R$, better approximates the electrostatic screening effects of bulk solvent. This set of simulations produced more consistent results, without abnormal β -turn twists seen previously, and without side chain-main chain hydrogen bonding in the turns. As a rule, β -bellin, containing L-Pro, tended to unfold. For example, the root-mean-square deviation between the starting, energy-minimized structure and the final "quenched" structure was 0.5 nm after 10 ps at 500° K, compared to the r.m.s. deviation of only 0.3 nm for the Gly- and D-Pro- β -bellins under the same conditions. The significantly higher stability of the Gly- and D-Pro β -bellins correlated with the opposite chiral tendencies of their respective β -turns. Thus, L-Pro imposed right-handed twists on turns (+23° average value), while Gly or D-Pro imposed left-handed twists (averages of -1° and -6°, respectively). This set of molecular dynamics simulations appeared to establish a direct role of β -turn twist (and, implicitly, side-chain chirality of the turns) in β -sheet stability.

(8) Criteria that discriminate between native proteins and incorrectly folded models.

Various theoretical concepts such as free energy potentials, electrostatic interaction potentials, atomic packing, solvent-exposed surface, surface charge distribution etc., were tested for their ability to discriminate between native proteins and misfolded protein models. Misfolded models were constructed by introducing incorrect side chains onto polypeptide backbones: side chains of the α -helical hemerythrin were modeled on the β -sheeted backbone of immunoglobulin VL domain while those of the VL domain were similarly modeled on the hemerythrin backbone. CONGEN, a conformational space sampling program, was

used to construct the side chains. Capability of the conformational search procedure to reproduce native conformations was gauged first by rebuilding (the correct) side chains in hemerythrin and the VL domain: constructs with r.m.s. differences from the X-ray side chains 2.2 - 2.4 Å were produced, and many calculated conformations matched the native ones quite well. Incorrectly folded models were then constructed using the same conformational protocol applied to incorrect amino acid sequences. All CONGEN constructs, both correctly and incorrectly folded, were characterized by exceptionally small molecular surfaces and low potential energies. Surface charge density, atomic packing and Coulomb formula-based electrostatic interactions of the misfolded structures and the correctly folded proteins were similar, and therefore of little interest for diagnosing incorrect folds. The following criteria clearly favored the native structures over the misfolded ones: (i) solvent-exposed side chain non-polar surface, (ii) number of buried ionizable groups, and (iii) empirical free energy functions that incorporate solvent effects.

List of Major Publications

J. Novotny & R. E. Bruccoleri: Correlation among Sites of Limited Proteolysis, Enzyme Accessibility and Segmental Mobility. *FEBS Letters* 211, 185-189 (1987).

J Allen, J. Novotny, J. Martin & G. Heinrich: Molecular Structure of Mammalian Neuropeptide Y: Analysis by Molecular Cloning and Computer-Aided Comparison with Crystal Structure of Avian Homologue. *Proc. Nat. Acad. Sci. U.S.A.* 84, 2532-2536 (1987).

J. Novotny, A. Rashin, R. E. Bruccoleri: Criteria that discriminate between native proteins and incorrectly folded models. *Proteins* 4, 19-30 (1988).

G. Evin, F.X. Galen, W. D. Carlson, M. Handschumacher, J. Novotny, J. Bouhnik, J. Menard, P. Corvol, E. Haber: Characterization of five epitopes of human renin from a computer model. *Biochemistry* 27, 156-165 (1988)

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., Oppermann, H. Protein engineering of antibody binding sites: Recovery of specific activity in an antidigoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Nat. Acad. Sci. U.S.A.* 85, 5879-5883 (1988)

Bruccoleri, R.E., Haber, E., Novotny, J. Structure of antibody hypervariable loops reproduced by a conformational search algorithm. *Nature* 335, 564-568 (1988)

Kraus, J.P., Novotny, J., Kalousek, F., Swaroop, M., Rosenberg, L.E. Different structures in the amino-terminal domain of the ornithine transcarbamylase leader peptide are involved in mitochondrial import and carboxyl-terminal cleavage. *Proc. Nat. Acad. Sci. U.S.A.* 85, 8905-8909 (1988)

List of Major Presentations

78th Annual Meeting of the American Society of Biological Chemists, Philadelphia PA (June 1987): "Computer Analysis of a Single-Residue Mutation that Affects Antibody Specificity (a joint poster presentation with R. E. Bruccoleri and E. Haber).

72nd Annual Meeting of the Federation of American Societies for Experimental Biology, Las Vegas, Nevada (May 1988): "Antibody combining site structure reproduced by conformational search algorithm".

1988 Sanibel Symposia on Quantum Chemistry, Marineland, Florida (March 1988): "Attribution of binding energy in antigen-antibody complexes McPC 603, D1.3 and HyHEL-5".

1988 Gordon Conference of Diffraction Methods in Molecular Biology, Salve Regina College, Newport Rhode Island (July 1988): "Attribution of binding energy in antigen-antibody complexes McPC 603, D1.3 and HyHEL-5".

1989 Joint Meeting of the American Society for Chemistry & Molecular Biology and The American Society for Cell Biology, San Francisco, California (January 1989): "Molecular dynamics simulations to explore twists of two-residue β -turns" (with J. & D. Richardsons)

Fig. 1

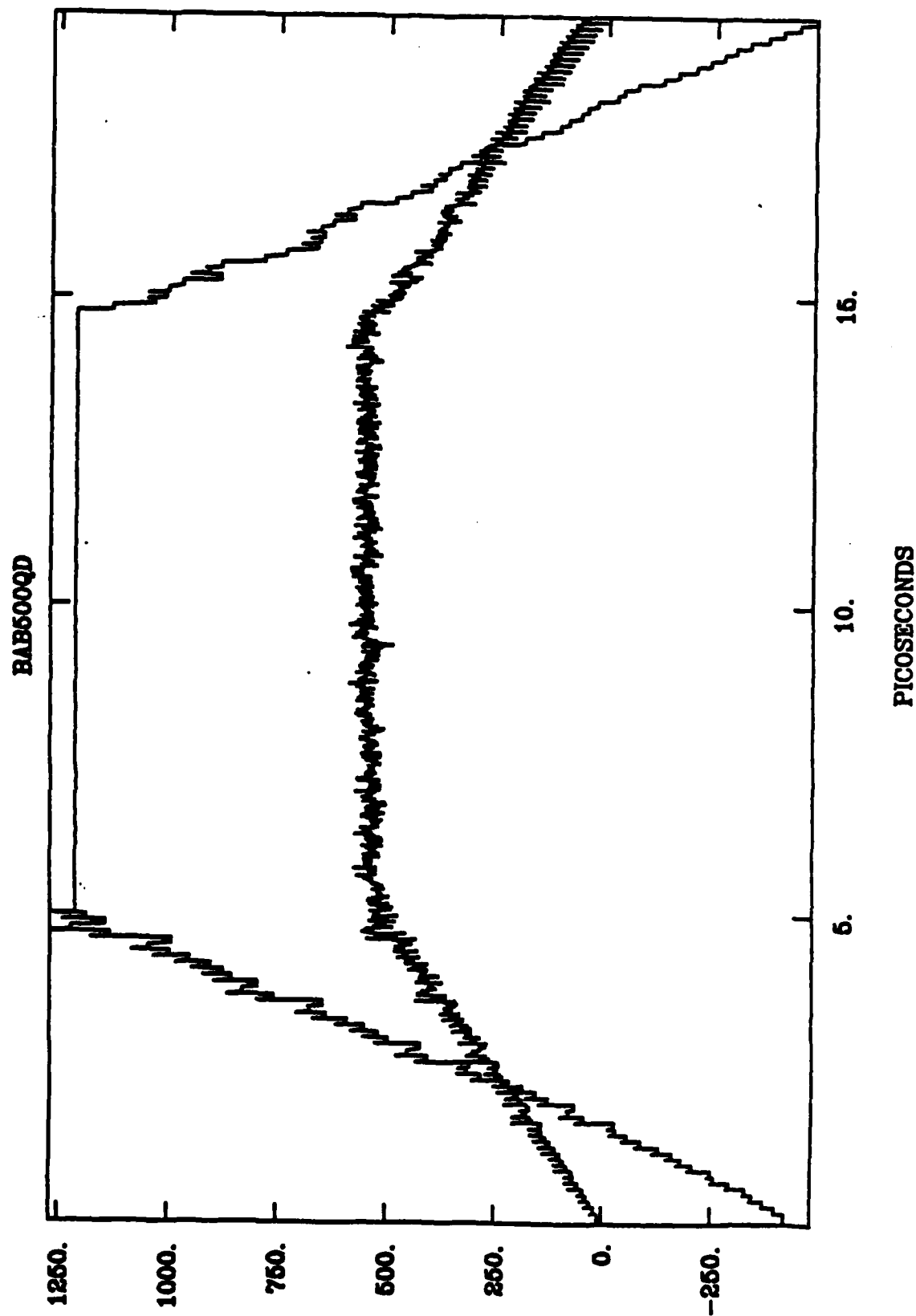
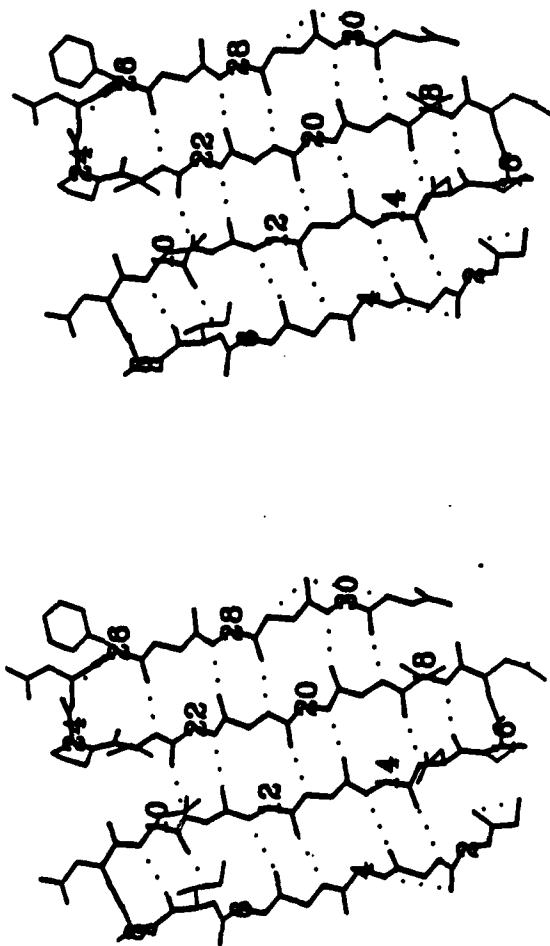
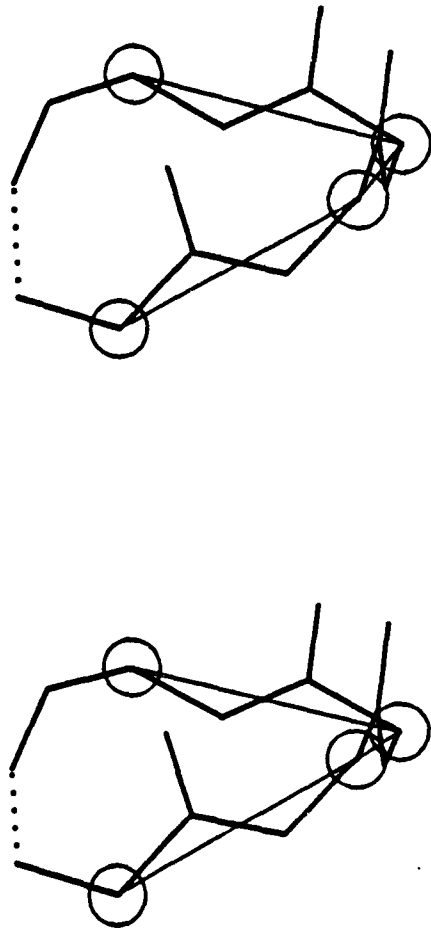


fig. 2



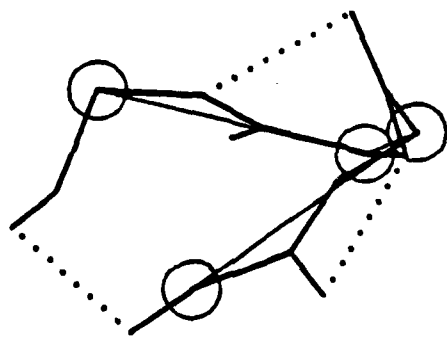
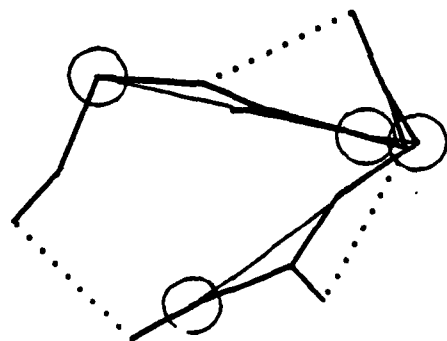
β -bellin chain B2, energy-minimized

Fig. 3



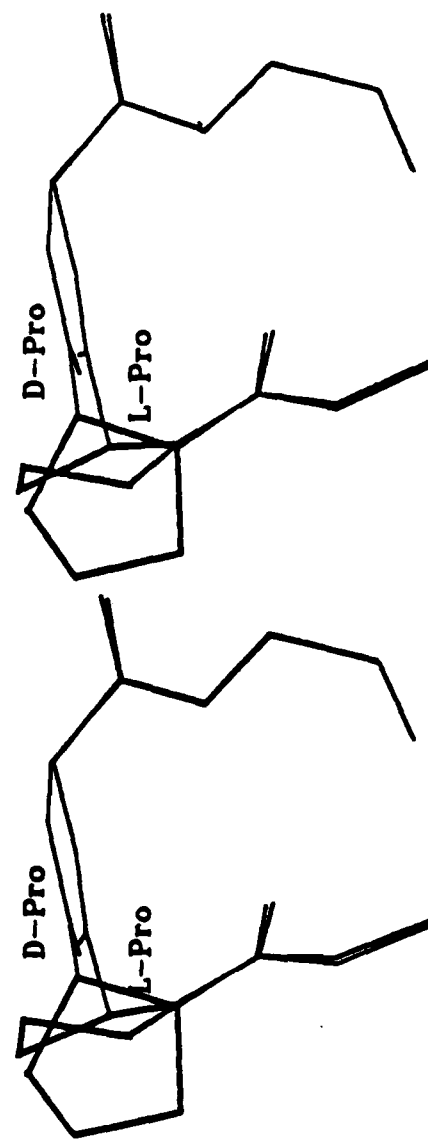
Loop 7-8-9-10 in β -bellin B2 chain: type I (Π') turn ($\text{Ca}-\text{Ca}-\text{Ca}-\text{Ca} = +45^\circ$)

fig. 4

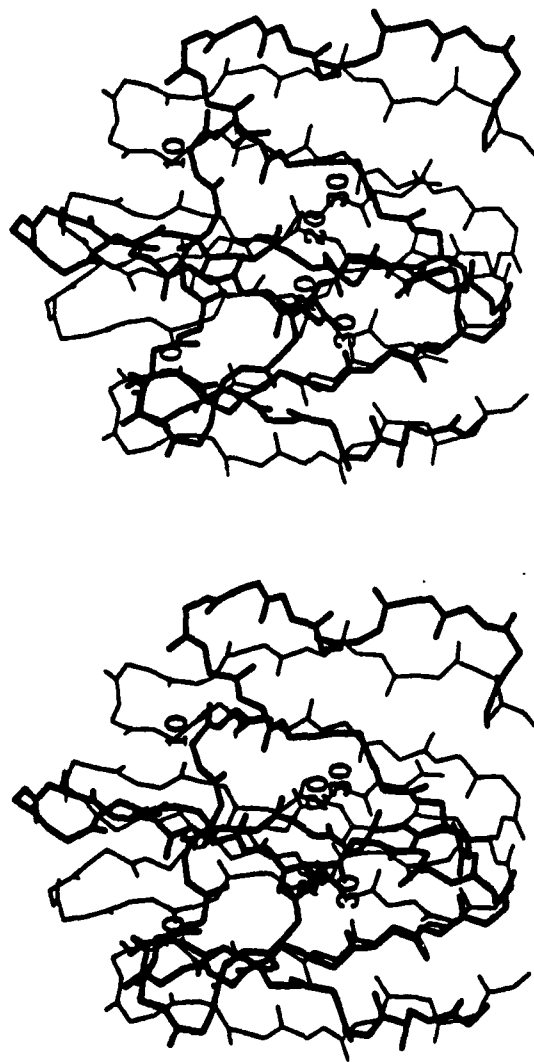


Loop 23-24-25-26 in β -bellin B2 chain: type I' (II) turn ($\text{Ca}-\text{Ca}-\text{Ca}-\text{Ca} = -45^\circ$)

Fig. 5



Loop Asp-Pro-Asn-Thr (15-18) in β -bellin: comparison of L-Pro 16 and D-Pro 16



β -bellin minimized and after 500 K dynamics, r.m.s. 5.0 Å
Implicit solvent by AMBER/CHARMM 19 parametrization & $T=4r'$ elec.

Fig. 1

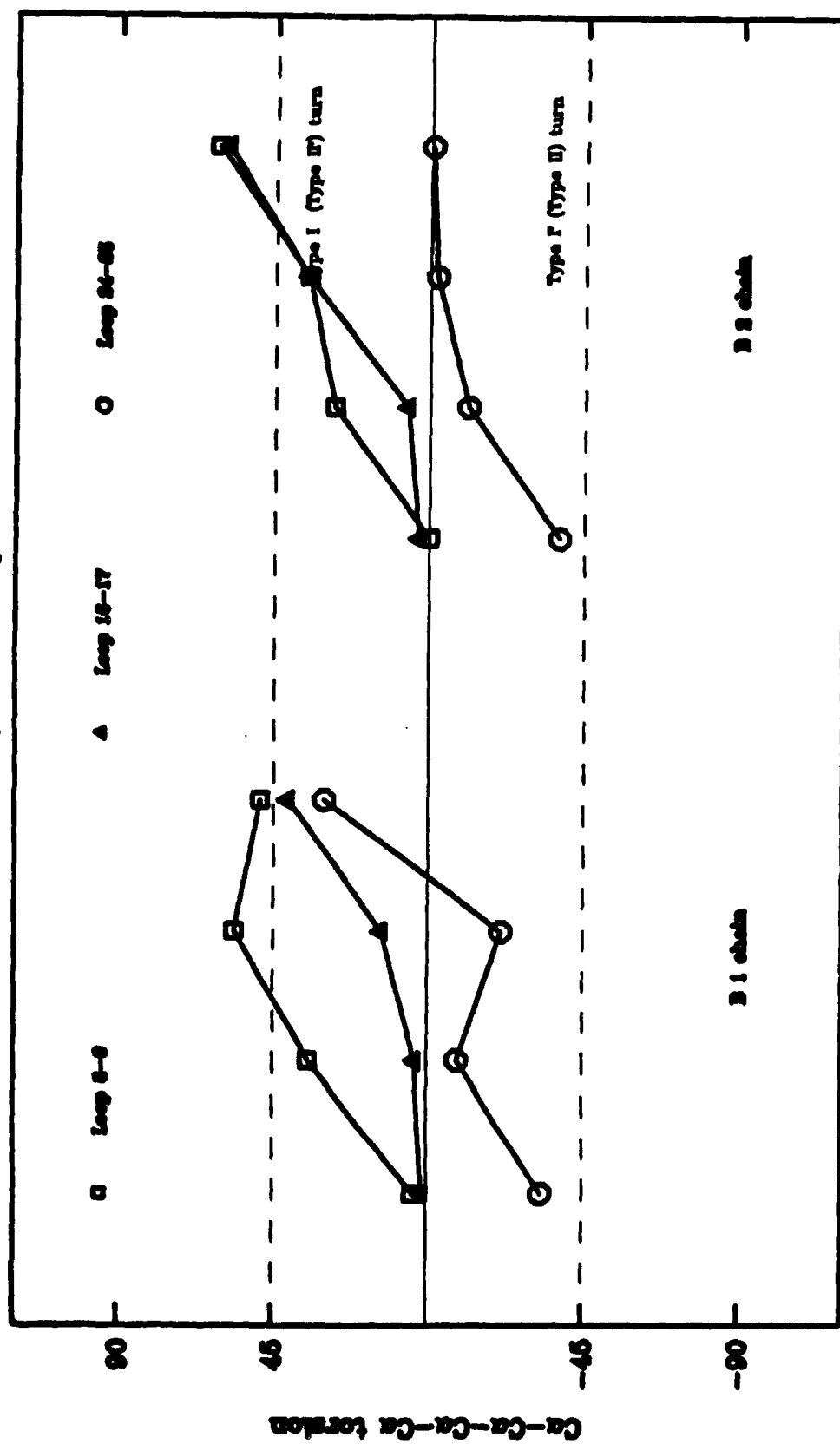


Gly- β -bellin minimized and after 500 K dynamics, r.m.s. 3.1 Å
Implicit solvent by AMBER/CHARMM 19 parametrization & $\epsilon=4\epsilon'$ elec.

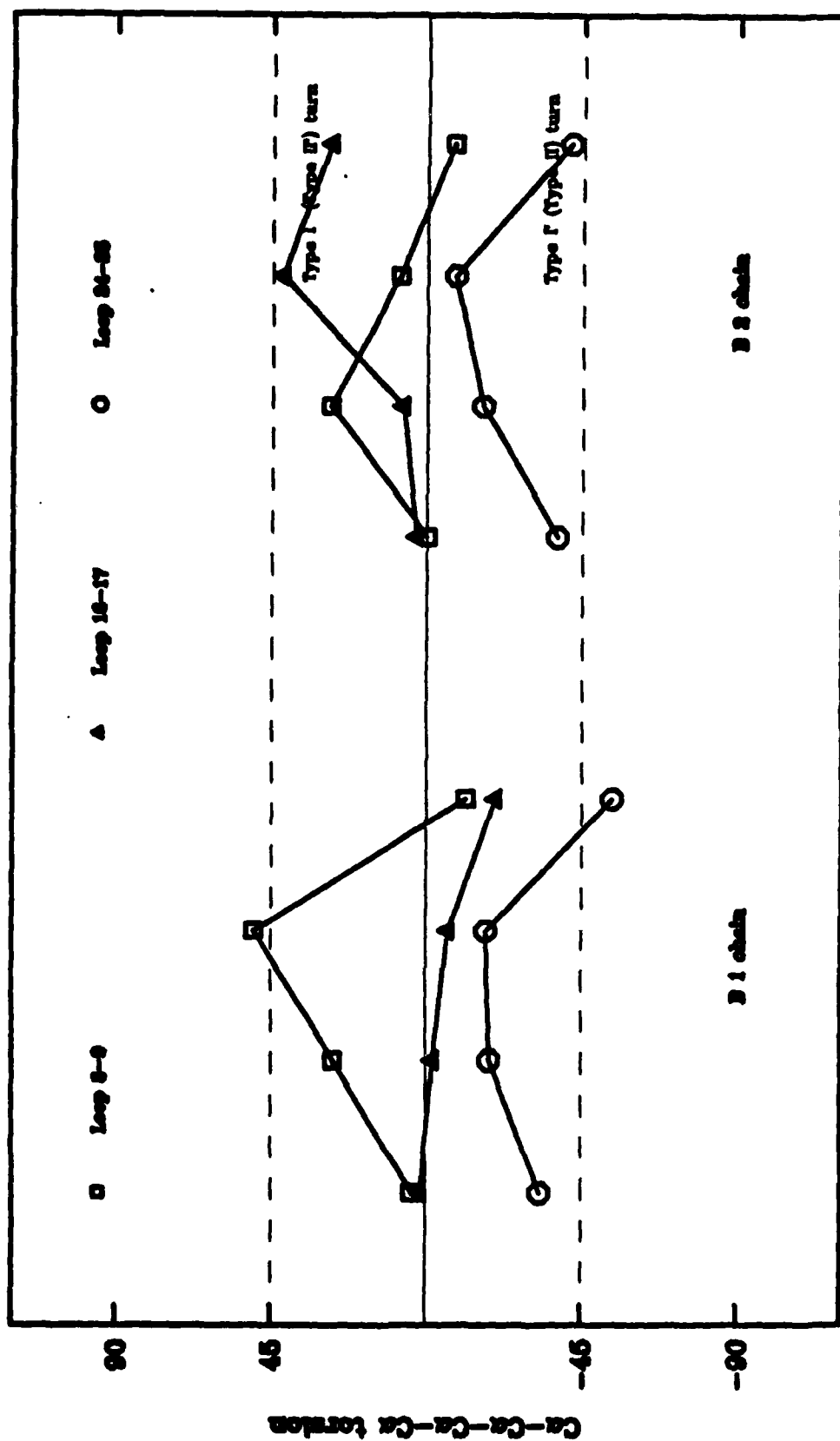
Explanations to graphs.

The three following graphs summarize β -turn twists (i.e., the Ca-Ca-Ca-Ca dihedral angles of the 4 successive residues of the two-residue turns) upon energy minimization and molecular dynamics simulations. The display contains, from left to right, progressive changes of turn dihedral angles as the original β -ballin coordinates (leftmost points) are energy minimized and subjected to dynamics at 300° K (the middle points) or at 500° K (the rightmost points). Turn dihedral angles are recorded separately for the two different β -ballin polypeptide chains (B1, B2) and the three turns of the same chain are distinguished by symbols: square (turn 7-8-9-10), triangle (turn 15-16-17-18) and circle (turn 23-24-25-26).

BAB- β -Bollin, loops



Gly-BAB- β -Ballin, loops



D-Pro-BAB- β -Ballo, loops

